UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR

Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

Análisis de las concurrencias de flujos en Internet

Víctor Martín Hernández
Tutor: José Luis García Dorado
Ponente: Javier Aracil Rico

JUNIO 2017

Análisis de las concurrencias de flujos en Internet

AUTOR: Víctor Martín Hernández
TUTOR: José Luis García Dorado

Hight Performance Computing and Networking research group
Dpto. de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Junio de 2017

Resumen

Han pasado ya más de 20 años desde que NetFlow de Cisco fuese patentado en el año 1996. Años en los que la monitorización y análisis de redes de comunicaciones basados en flujos de red ha ganado extraordinaria relevancia en tareas fundamentales como asegurar las más alta calidad de servicio (QoS) posible, el dimensionado de la capacidad de equipos hardware o la identificación de vulnerabilidades como por ejemplo ataques de denegación de servicio (DoS). En este trabajo se pretende aportar en la dirección de un mejor conocimiento de las dinámicas de los flujos de red en redes comerciales tanto longitudinalmente (esto es, con el tiempo) como espacialmente (varios enlaces). En concreto, estudiamos la relación del número de flujos activos (o concurrentes) frente al ancho de banda medido, que denominamos ratio de flujos. Esta métrica es fundamental para un gestor de red que si bien conoce el ancho de banda esperado en su red carece típicamente de ninguna intuición de cuál va a ser la carga de aquellas aplicaciones basadas en flujos de red. Este trabajo analiza esta métrica en múltiples trazas disponibles en la institución CAIDA, concluyendo que en escenarios habituales el ratio de flujos es menor a 300 flujos activos por Mb/s y con cierta normalidad hasta los 400. Mientras que en situaciones anómalas se han medido hasta 1200 flujos por Mb/s. Por el contrario, el estudio de la homogeneidad de la métrica durante 7 años y por sentidos de enlaces entre ciudades como Chicago, Seattle, San José y Los Ángeles ha mostrado diferencias significativas debidas probablemente a cambios en los servicios transportados por estos enlaces.

Abstract

Since Cisco’s Netflow was patented in 1996 has become a relevant tool in networks tasks such as ensuring the highest quality of service (QoS), the planning of hardware equipment in terms of performance, and the identification of vulnerabilities (for example, Denial-of-service attacks (DoS)). In this work, we aim at contributing in the further knowledge of the dynamics of such Netflows in commercial networks both temporally (i.e., over time) and spatially (several links). Specifically, we study the ratio between the number of active (or concurrent) flows and the average bandwidth, namely ratio of flows. This metric is key for networks managers as they usually know the expected bandwidth in their network but they tend to ignore how this is going to translate into burden for those application based on Netflows. This work analyses such metric in several available traces by CAIDA institution concluding that in regular scenarios the ratio of flows is below 300 active flows per Mb/s and, still inside normality, up to 400 ones. While in anomalous scenarios, it has been measured up to 1200 flows per Mb/s. Moreover, the study of homogeneity of the metric during 7 years and per direction between cities such as Chicago, Seattle, San Jose and Los Angeles has shown significant differences likely because of changes on the services carried by such links.

Palabras clave

Monitorización de redes, NetFlow, flujo de red, ratio de flujos, CAIDA.

Keywords

Network monitoring, NetFlow, ratio of flows, CAIDA.

Agradecimientos

A mis padres y hermano.

A aquellos que, desde pequeños me han acompañado a lo largo de toda mi enseñanza, los chavales: Pablo, Alex, Raúl y Adri. A todos mis compañeros que me han apoyado durante el paso por la Universidad, especialmente a mis compañeros Alex, Ricar, Santi, Cifu, Marino y, mi inseparable compañero de prácticas, Miguel. A mi tutor, José Luis García Dorado, por toda la atención y ayuda que ha sido capaz de proporcionarme y que, sin duda, no habría sido posible sin él.

Gracias por estar ahí.

INDICE DE CONTENIDOS

INDICE DE FIGURAS

INDICE DE TABLAS

1 Introducción

1.1 Motivación

A lo largo de los años, el uso de la Internet ha crecido enormemente tanto en usuarios domésticos como en las pequeñas, medianas y, sobre todo, grandes empresas.

Esto ha desembocado en un gran interés por la monitorización del tráfico dado que es una de las herramientas fundamentales para proveer calidad de servicio (QoS) adecuada a los usuarios, así como mantener los costes de despliegues y sus actualizaciones en cifras razonables. Por ejemplo, el análisis de los datos de la red permitirá, a las empresas proveedoras de servicios en la Internet (ISP), analizar de una forma eficaz aquellas necesidades tanto actuales como futuras de las redes que manejan. De esta forma, serán capaces de tomar decisiones más acertadas sobre qué tecnologías, arquitecturas y equipos deben utilizar, con el fin de maximizar el QoS que proporciona a los clientes.

De manera similar muchas empresas deben invertir grandes cantidades de dinero en campos relacionados con la seguridad informática. Esto se debe a la proliferación de diferentes tipos de vulnerabilidades, destacando especialmente los ataques de denegación de servicios (DoS) o los ataques de denegación de servicio dirigidos (DDoS), cuyo objetivo es la sobrecarga de los enlaces de red a través de la ocupación completa de éstos mediante el envío indiscriminado de paquetes. De manera que, la detección temprana de estos tipos de ataques será esencial para minimizar los daños que puedan ser ocasionados. Para ello, una vez más, podremos emplear las tecnologías de monitorización y análisis del tráfico de datos en las redes de comunicación basados en flujos concurrentes.

Dentro de las herramientas que los gestores de red disponen para ello, destaca los registros de flujos de red. Un flujo de red se define como un conjunto consecutivo de paquetes que pasan por un punto durante un intervalo de tiempo (típicamente, 15 segundos como *timeout* de actividad) y comparten cinco (en ocasiones siete) atributos. Estos atributos serán: dirección origen y destino IP, puerto origen y destino y la capa de transporte (añadiéndose la clase de servicio y la interfaz del *router* o *switch* en caso de considerar los siete atributos).

Y es que debido al aumento de las velocidades de transmisión en las redes actuales, una monitorización y análisis a nivel de paquete se ha convertido en una opción inviable por múltiples motivos: procesamiento lento, velocidad de almacenamiento alta, sobrecarga de la red por la exportación de grandes tablas de estadísticas, etc. En el otro extremo la monitorización basada en estadísticas agregadas (por ejemplo, series temporales de ancho de banda o de uso de puertos) provee una información, que si bien útil para identificar que hay un problema, es muy limitada para concretar las causas raíces del mismo [1]. Es por ello que muchos sistemas se basan en la creación de flujos de red, que son posteriormente explotados por múltiples aplicaciones.

Algunas de estas aplicaciones son:

- Estudio de las estadísticas generadas en el análisis para el desarrollo de nuevas soluciones y modelos de red.

- La detección de tráfico no autorizado o malicioso con el fin de evitar modificaciones costosas, identificando los usuarios o conexiones que causan la congestión.

- Almacenaje de datos comprimidos/resumidos que se encuentra en forma de paquetes para traducirlos en un conjunto de flujos. Éstos proveerán una información basada en la conexión punto a punto realmente valiosa, llegando a utilizar únicamente el 5% de los datos proporcionados por los paquetes [5].

- Sistemas de identificación y clasificación de tráfico [2].

- Sistemas de control de calidad de telefonía VoIP [3].

- Caracterización de tráfico y modelado.

De hecho existen múltiples sistemas en operación de monitorización basada en flujos como NetTraMet [6], Ntop [7], NG-MON [8], o todas aquellas variantes basadas en NetFlow [9], estos últimos estandarizados por el IETF, Internet Engineering Task Force, a través de IPFIX [10].

En conclusión existen variadas aplicaciones que necesitan la construcción de flujos para su correcta operación.

1.2 Objetivos

El presente Trabajo de Fin de Grado tiene como objetivo principal dar una respuesta clara a la pregunta de qué carga puede esperar una aplicación basada en flujos de red dado un ancho de banda real o esperado en una red de comunicaciones. En otras palabras que número de flujos concurrentes se puede prever que transporta una red conocida su capacidad o uso en términos de ancho de banda. Para ello se recurrirá a trazas tan generales y durante el máximo tiempo que se pueda.

De este modo ante el despliegue de una aplicación que use flujos de red, y conocido el ancho de banda se podrá configurar y dimensionar la capacidad del hardware donde corre tal aplicación. Este hardware incluye la capacidad de la CPU, de disco duros, tasa I/O de distintos puerto (por ejemplo, PCI Express), entre otros.

Además, al principal objetivo se suman los siguientes:

- Familiarización con la monitorización y el análisis de la red previa elaboración de programas y obtención de resultados con el fin de abordar los problemas bajo el conocimiento adecuado.

- Desarrollo de una aplicación para el análisis capaz de procesar una gran cantidad de tráfico, debido a la gran cantidad de medidas y capturas que disponemos, y proporcionarnos la información necesaria para, posteriormente, obtener las estadísticas.

- Estudio del comportamiento de la red a través de la generación de estadísticas y gráficas partiendo de los informes proporcionados por el programa desarrollado anteriormente.

Contribuir con el presente documento a la comunidad de la Internet en su avance en un mejor conocimiento de las dinámicas de la Internet [4].

1.3 Fases de realización

Con el fin de llevar la tarea lo más ordenada posible y obtener unos resultados óptimos, ha sido necesario seguir una serie de fases:

- **Documentación.** Documentarse a cerca de los flujos de red, sus aplicaciones, sistemas de monitorización y análisis de tráfico, y demás características. Incluimos en esta fase el aprendizaje propio de Matlab y el uso de su entorno.

- **Obtención de medidas de red lo más generales posibles y durante un tiempo significativo.** Para ello, se descarga todas las trazas *pcap* disponibles en CAIDA, albergándolas en uno de los servidores facilitados por la Universidad Autónoma de Madrid y localizado en la Escuela Politécnica Superior.

- **Implementación del programa Simulador NetFlow.** Implementación del programa, en lenguaje de programación de alto nivel, C, para la generación de los informes que recopilan la información, por días, de las trazas *pcap*.

- **Generación de informes.** Generación de los informes que recopilan la información, por días, de las trazas *pcap* descargadas previamente.

- **Generación de estadísticas y elaboración de gráficas.** Generación de las estadísticas a partir de los informes generados desde el programa Simulador NetFlow, así como la elaboración de las gráficas para la obtención de resúmenes visuales.

- **Redacción de la memoria.** Redacción del documento que se presenta para poder plasmar el conocimiento y la experiencia adquirida en este Trabajo de Fin de Grado.

1.4 Estructura del documento

La memoria consta de seis capítulos, que son los siguientes:

- **Sección 1: Introducción.** La presente sección marca la motivación de la realización de este Trabajo de Fin de Grado, exponiendo los objetivos principales y el planteamiento ideado para su conclusión.

- **Sección 2: Estado del arte.** Evolución que ha sufrido NetFlow desde sus inicios hasta llegar a día de hoy, junto con la explicación del término flujo de red y el cambio sufrido en los sistemas de análisis y monitorización del tráfico. Además, una breve explicación sobre qué es CAIDA, sus objetivos y su expansión global.

- **Sección 3: Datos disponibles e implementación.** Características de los datos proporcionados por CAIDA, ligada a la explicación de la obtención de los informes generados a través de un programa cuya funcionalidad es hacer de *simulador Netflow*, implementado por nosotros mismos, así como una explicación de las estadísticas generadas por Matlab partiendo de los informes.

- **Sección 4: Análisis de resultados.** Estudio de las estadísticas y gráficas generadas previamente a través de los diferentes *scripts* desarrollados en Matlab, localizando las zonas que contienen un número de flujos normal, frente a aquellas que consideraremos anómalas, así como un estudio espacial a lo largo de los años y un estudio longitudinal de los diferentes enlaces disponibles, y sus distintas direcciones que los componen.

- **Sección 5: Conclusiones.** Se plasmarán las conclusiones finales a tomar tras la realización de este Trabajo de Fin de Grado.

- **Sección 6: Trabajo futuro.** Se proporcionará una breve lista de posibles trabajos futuros a realizar para ampliar el análisis de los flujos de red.

2 Estado del arte

2.1 Introducción

Esta sección muestra el desarrollo previo de la evolución de NetFlow, así como una breve explicación sobre la organización que proporciona trazas públicas, CAIDA.

En primer lugar, hablaremos sobre la evolución que ha sufrido NetFlow desde sus inicios hasta llegar a día de hoy, la explicación de que es un flujo de red, el progreso de los sistemas de monitorización y análisis de tráfico de red, y el proceso NetFlow. Seguidamente, comentaremos las diferentes aplicaciones que se pueden dar, a día de hoy, los flujos de red y como han ido creciendo en los últimos años y, para finalizar, haremos una introducción a CAIDA, explicando sus objetivos y cómo está distribuido a lo largo del mundo.

2.2 NetFlow

Originalmente, la tecnología NetFlow fue diseñada por Cisco con un único fin, su introducción en los *routers* y *switches* Cisco para la más rápida conmutación de paquetes. Dicha tecnología fue desarrollada en el año 1996 por Darren y Brarry Bruins para el propio software Cisco y fue introducida a partir de la versión IOS 11.x.

En un primer momento, Cisco Systems implantó IOS NetFlow en los *routers* y *switches* Cisco para generar dinámicamente unas tablas de enrutamiento a medida que llegasen los paquetes. De esta forma, cada uno de éstos, pertenecientes a un mismo flujo, serían enrutados rápidamente sin necesidad de inspeccionar extensamente sus características y, con ello, reducir considerablemente el tiempo de retención del paquete en el interior del *router* o *switch*. Por tanto, IOS NetFlow no sería otra cosa más que una *caché* IP, tal y como vemos en la Figura 2-1.

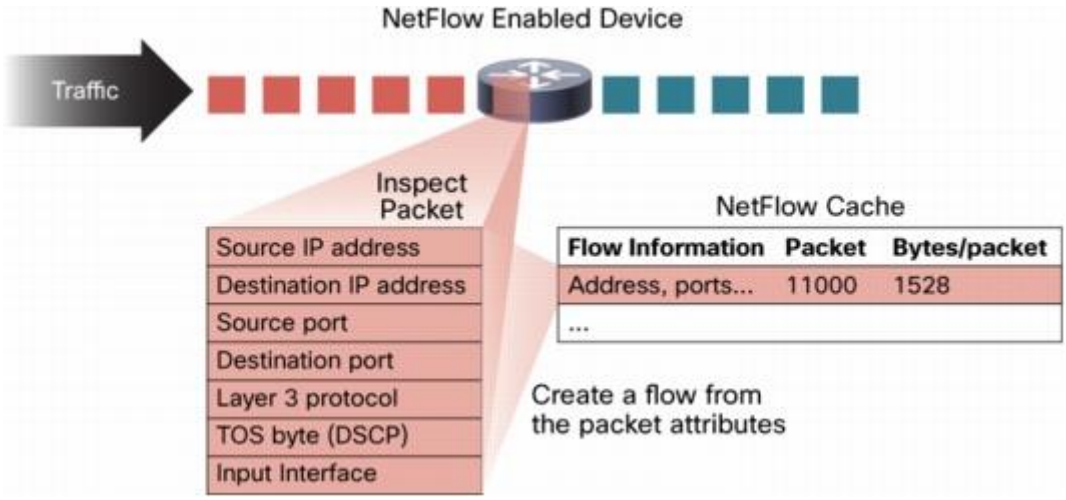


Figura 2-1: Creación de una entrada en la caché NetFlow [9]

IOS NetFlow avanzaba y, pocos años más tarde, se observó que la información que recopilaba en las tablas de flujos activos generadas era más valiosa de lo que en un principio se creía. No solo se recopilaba la información de conexión punto a punto más destacada, sino que tan solo se necesitaría un 5% de todo el conglomerado de datos proporcionado por los paquetes [5].

Tradicionalmente, un flujo de red, construido a nivel IP, se define como un conjunto consecutivo de paquetes que pasan por un punto durante un intervalo de tiempo (típicamente, 15 segundos como *timeout* de actividad) y comparten cinco (en ocasiones siete) atributos. Estos atributos serán: dirección origen y destino IP, puerto origen y destino y la capa de transporte (añadiéndose la clase de servicio y la interfaz del *router* o *switch* en caso de considerar los siete atributos).

El porqué de este conjunto de atributos es claro, con ellos tendremos la mayor información posible acerca de un conjunto de paquetes, tales como: quién origina el tráfico, quién lo recibe, qué aplicación lo está utilizando (orientativamente, a partir de los puertos) o la prioridad de éste (establecida por la clase de servicio), entre otros. Adicionalmente, puede acompañarse de otro tipo de información en caso de añadir más atributos a cada entrada de la tabla *caché* de flujos activos.

Por tanto, para la obtención de toda esta información, se emplean sistemas de monitorización y análisis de tráfico en tiempo real. Inicialmente, dichos sistemas se componían de arquitecturas con un único elemento que realizaba todas las acciones de observación y estudio de datos en las redes de comunicaciones. Algunos ejemplos de éstos son las primeras versiones de Ntop [7] y argus.

Posteriormente, las arquitecturas de estos sistemas fueron evolucionando junto con la necesidad de incrementar la velocidad en los enlaces. De esta forma, éstas se integraron por diferentes componentes agrupados y trabajando en paralelo, siguiendo el estándar de IETF, Internet Engineering Task Force, a través de IPFIX [6]. El fin de esto es obvio, permitir la escalabilidad y flexibilidad de los sistemas, así como minimizar la sobrecarga de los sistemas.

Estas arquitecturas están conformadas comúnmente por tres componentes de procesamiento: generador de flujos, contenedor de flujos y analizador del tráfico. El generador de flujo será el encargado de capturar los paquetes mediante el uso de funciones de *mirroring* o *splitting* para, posteriormente, exportarlo al contenedor de flujos. También será posible exportar los datos directamente desde los *routers* o *switches Cisco* a éste. El contenedor de flujos almacenará los datos, hasta tener tantos como se consideren oportunos, antes de ser enviados al analizador.

En la Figura 2-2 y la Figura 2-3 podremos observar de una forma más clara el proceso que sigue el flujo, desde la creación de la entrada en la tabla *caché* hasta su análisis, así como los componentes que componen los sistemas de monitorización y análisis de flujos.

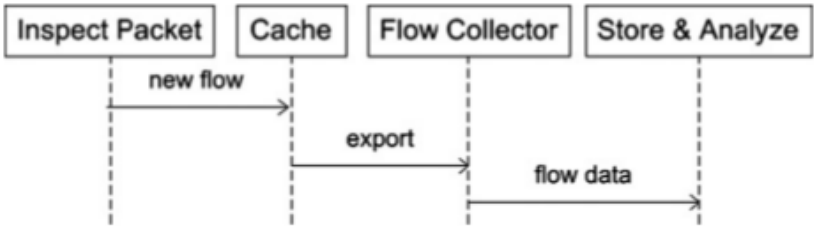


Figura 2-2: Proceso NetFlow [11]

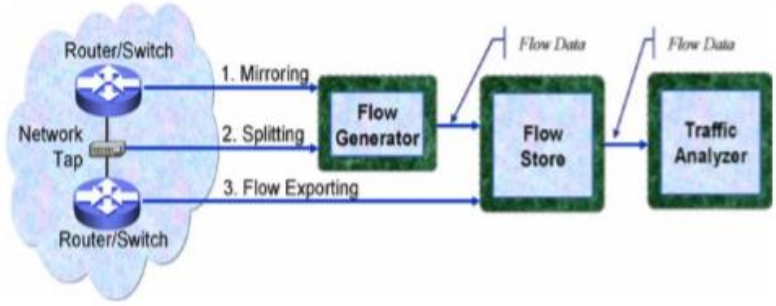


Figura 2-3: Arquitectura en paralelo NetFlow [12]

2.2.1 Perspectivas NetFlow

A continuación, vamos a comentar las principales perspectivas de investigación de las aplicaciones basadas en flujos de red.

La monitorización y el análisis de la red proporcionan información valiosa a los administradores de red, los proveedores de servicios de la Internet y proveedores de contenido. Comparando con otras tecnologías, como SNMP o *Windows Management Instrumentation* (WMI), los datos de flujo de la red contienen información adicional para un análisis posterior. La monitorización basada en NetFlow puede ser categorizada como:

- Monitorización de red. Proporciona información acerca de los *routers* y *switches*, así como una vista base de toda la red, y se utiliza para la detección de problemas y para dar una solución eficiente a dicho problema.

- Monitorización de aplicaciones. Proporciona información sobre el uso de aplicaciones a través de la red, y se utiliza para la planificación y asignación de recursos.

- Monitorización de *host*. Proporciona información sobre el modo de empleo que el usuario realiza relativo a las redes y aplicaciones, y se utiliza para la planificación o el control de acceso a la red.

- Monitorización orientada a la seguridad. Proporciona información sobre los cambios de comportamiento de la red, y se utiliza para identificar ataques DoS, virus y gusanos, y anomalías de red.

- Contabilidad y facturación. Proporciona la medición de red y se utiliza para facturación.

De este modo, podemos observar en la Figura 2-4, aunque algo desactualizada, pero con un comportamiento similar en los últimos años, como las perspectivas de monitorización y seguridad han ido creciendo de una manera casi exponencial en los últimos años.

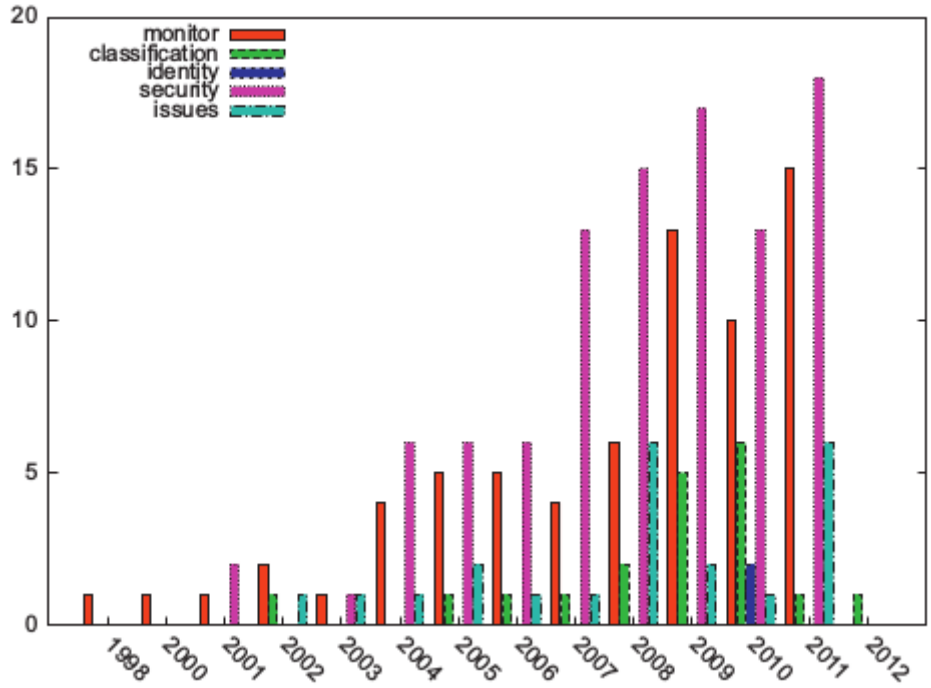


Figura 2-4: Perspectivas analizadas por años [11]

2.3. CAIDA

El Centro para el Análisis de Datos Aplicados de Internet, CAIDA, en inglés Center for Applied Internet Data Analysis, se encuentra en la sede federal de SDSC, en inglés San Diego Supercomputer Center, ubicado en la Universidad de California, San Diego. El centro está compuesto por una colaboración de entidades gubernamentales, de investigación y comerciales, que trabajan juntas para mejorar la Internet.

CAIDA realiza la investigación de la red y la construcción de infraestructuras de investigación, con el fin de apoyar la recolección de datos a gran escala, la conservación y la distribución de éstos para su posterior investigación por parte de la comunidad científica. Por tanto, tiene como principales objetivos:

- Proporcionar una visión de la función macroscópica de la infraestructura de la Internet, el comportamiento, uso y evolución.

- Fomentar un entorno de colaboración en el que los datos puedan ser adquiridos, analizados y compartidos.

- Mejorar la integridad del campo de la ciencia de la Internet.

- Informar a la ciencia, la tecnología y políticas públicas de comunicaciones.

Además, CAIDA despliega y mantiene una plataforma de medición distribuida a nivel mundial que ha ido creciendo enormemente a lo largo de los años. Un ejemplo actual de ello lo tenemos en la Figura 2-3.

Figura 2-5: Mapa con la localización de las plataformas de medición de CAIDA [13]

3 Datos disponibles e implementación

3.1 Introducción

En esta sección comentaremos el proceso llevado a cabo para la caracterización de los datos sometidos a estudio mediante *Netflows*.

Comenzaremos explicando las características de los datos descargados, que serán utilizados posteriormente para obtener ciertos datos a través de un programa cuya funcionalidad es hacer de simulador *Netflow*, implementado por nosotros mismos y, que será lo siguiente sobre lo que hablaremos. Para finalizar, comentaremos la herramienta empleada en el análisis de las estadísticas proporcionadas por dicho programa.

3.2 Datos

La colección de trazas han sido obtenidas en su totalidad del portal de CAIDA [14], correspondientes a dos enlaces localizados en Estados Unidos.

El primero de ellos se ubica en el estado de Illinois, más concretamente en el centro de datos Equinix [15] de Chicago, y está conectado a un enlace de la red troncal de un ISP Tier1 entre Chicago y Seattle, Washington.

El segundo se ubica en el estado de California, en el centro de datos Equinix de San José, y está conectado a un enlace de la red troncal de un ISP Tier1 entre San José y Los Ángeles, California.

Generalmente, las trazas están divididas de acuerdo a diferentes parámetros, como son: año, enlace al que pertenecen, día y mes del año, y la dirección. Con dirección nos referimos tanto a la de ida como de vuelta, esto es, en el caso del primer enlace visto, tendrá una dirección de Chicago a Seattle y otra de Seattle a Chicago.

Todas las trazas fueron tomadas entre la 13:00 y 19:00 de la tarde, cada una de ellas con una duración aproximada de una hora, divididas en múltiples *pcap* y tomadas minuto a minuto, esto es, un día concreto estará formado por múltiples trazas de un minuto.

Al tratarse de las horas con más tráfico en la red, obtendremos un análisis sobre escenarios de mayor carga (vea por ejemplo la tráfico en AMS-IX (Amsterdam Internet Exchange) en la Figura 3-1), típicamente más útiles para el dimensionado [16].

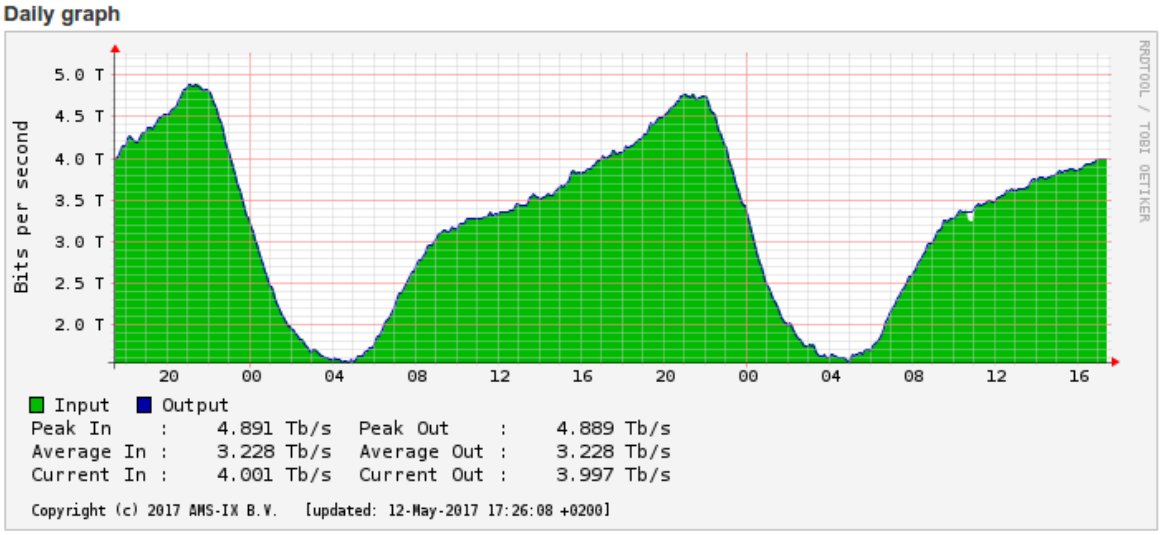


Figura 3-1: Gráfico del tráfico de red por horas en AMS-IX

Es importante saber de cara a un análisis posterior que, con el fin de anonimizar las trazas de tráfico del conjunto de datos, CAIDA ha empleado el prefijo de preservación del anonimato CryptoPan [17]. Además, se ha eliminado la carga útil de todos los paquetes.

Para finalizar, queremos destacar la dificultad añadida con la que hemos tenido que lidiar en cuanto a la obtención de tanta cantidad de información y el trabajo con ésta se refiere, llegando a tener que almacenar los datos en diferentes discos duros debido a la proporción de tamaño que ocupaba, estando en una escala de varias decenas de Terabytes.

3.2.1 Datos disponibles

Los ficheros *pcap* que hemos utilizado en este trabajo comienzan el 19 de marzo de 2008 a las 19.00h (GMT) para el enlace de Chicago, y el 17 de julio de 2008 a las 13.00h (GMT) para el enlace de San José. Concluirán el 17 de diciembre de 2015 a las 13.00h (GMT) para el enlace de Chicago, y el 19 de junio de 2014 para el enlace de San José. Sin embargo, se han encontrado algunas limitaciones en cuanto a la disponibilidad de los datos que conducen al descarte de varios puntos de presencia en el análisis.

Con el fin de tener una visión sobre los datos totales disponibles, se ha realizado una tabla en la cual se muestra al detalle los archivos *pcap* disponibles, a modo de ejemplo la Tabla 3-1 muestra los datos para 2008. La disponibilidad vendrá dada por un guión (-) cuando haya una ausencia de datos para el día del año marcado, o con una equis (X) en caso contrario. El resto de trazas disponibles se muestra en el Anexo A.

Año	Enlace	Mes	Dirección	Disponibilidad
2008	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	-
			Seattle-Chicago	-
		03	Chicago-Seattle	-
			Seattle-Chicago	X
		04	Chicago-Seattle	X
			Seattle-Chicago	X

Tabla 3-1: Perspectiva de los archivos disponibles (ver Anexo A)

3.3 Implementación y herramientas empleadas

El proceso de recolección de datos comienza por la creación de un programa capaz de simular el funcionamiento de un sistema de Netflow, esto es, la tabla de flujos activos a grano de 1 segundo. Tras obtener estos valores, se extraerán estadísticas a través de otro software externo que nos permita realizar los cálculos y gráficas que nos serán útiles para una correcta explicación, y sobre todo, llegar a conclusiones.

3.3.1 Simulador NetFlow

Nuestro simulador NetFlow tendría que ser capaz de procesar archivos de datos *pcap* aplicando ciertos filtros que puedan ser introducidos a la hora de ejecutar el programa, los cuales veremos a continuación.

Cabe destacar que durante el desarrollo del programa se tuvo que tener en cuenta en todo momento la optimización de recursos sobre los que hacía uso éste, así como la rapidez de ejecución.

Prueba de ello, lo tenemos en la Figura 3-2, que se corresponde con una de las salidas del programa, en la cual se nos muestra la fecha de inicio, la fecha de finalización, el número de paquetes procesados y su distribución en los diferentes protocolos.

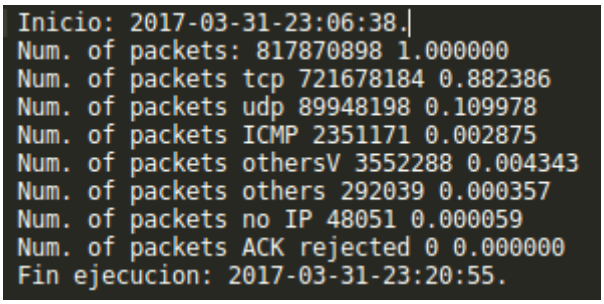


Figura 3-2: Ejemplo salida programa (fichero 1)

De esta manera, podemos concluir que nuestro programa será capaz de tratar 817 millones de paquetes en un total de 17 minutos o, lo que es lo mismo, será capaz de procesar 48 millones paquetes por minuto, y cerca de 3000 millones por segundo. Cifra por encima de una interface de 100 Gb/s, probando su alto rendimiento

Al ejecutar nuestro programa se nos pedirá ciertos argumentos de entrada como son: la ruta del fichero de entrada, el nombre del fichero de entrada, si se trata de un fichero *pcap* o de un fichero de texto plano, el tiempo de expiración de los flujos, la ruta de salida donde se guardará el informe generado y el tipo de protocolo (TCP, UDP, ninguno de ambos o todos).

La elección de poder introducir un archivo *pcap* o un fichero de texto plano es debido a la forma en que se organizan las trazas diarias proporcionadas por CAIDA. Como comentábamos en el punto anterior, un día está compuesto por múltiples ficheros *pcap* con una duración de un minuto. Por tanto, el objetivo era poder proporcionar la ruta en donde se encuentren los ficheros y una lista de todos ellos para que se ejecutasen seguidos y poder proporcionar un único informe diario.

De manera que, el programa hará uso de la librería *libpcap* con la que nos será más sencillo el trabajo de cara a la obtención e inspección de los paquetes de la traza. Como ya mencionábamos anteriormente, los datos a analizar forman parte de trazas capturadas previamente, por lo que se hará uso de la funcionalidad de apertura *offline*.

3.3.2 Salida del Simulador NetFlow

La salida de nuestro programa que simulará los registros obtenidos por NetFlow nos arrojará aquellos datos interesantes para su posterior uso en el estudio y análisis de estos mismos. Alguno de dichos datos serán los citados a continuación:

- ➔ Tiempo, que indica la fecha de llegada del paquete.
- ➔ Flujos activos (número de flujos activos en un instante dado).
- ➔ Ancho de banda.

Aunque reportamos otros datos como los bytes en ventana Ethernet, los bytes en ventana IP o el ancho de banda Ethernet consumido, no los hemos considerado importantes en nuestro posterior análisis. La Figura 3-3 muestra un ejemplo de esta salida.

num_muestras	timestamp(sec)	flujosactivos	bytesventana_ETHERNET	bytesventana_IP	totFlow	Mbps_ETHERNET	Mbps_IP
1	1205953148	74431	298739634	298529911	74431	2389.92	2388.24
2	1205953149	116016	323528771	323299122	116016	2588.23	2586.39
3	1205953150	146465	321585345	321349461	152298	2572.68	2570.80
4	1205953151	173486	335928303	335692136	185544	2687.43	2685.54
5	1205953152	198326	333764386	333536566	216535	2670.12	2668.29
6	1205953153	222278	330653398	330417783	246668	2645.23	2643.34
7	1205953154	245063	334069550	333844110	275566	2672.56	2670.75
8	1205953155	266572	336824429	336591599	303355	2694.60	2692.73
9	1205953156	287000	337442206	337214989	330309	2699.54	2697.72
10	1205953157	307364	339802094	339566204	356896	2718.42	2716.53
11	1205953158	327495	334733188	334495743	383133	2677.87	2675.97
12	1205953159	346756	334923328	334690053	408588	2679.39	2677.52
13	1205953160	365918	336656967	336433002	433673	2693.26	2691.46
14	1205953161	384982	336093695	335862906	458628	2688.75	2686.90
15	1205953162	403928	343488982	343260972	483464	2747.91	2746.09
16	1205953163	423136	340431818	340205898	508637	2723.45	2721.65
17	1205953164	423832	338635454	338403401	533363	2709.08	2707.23
18	1205953165	423502	332498686	332273383	558063	2659.99	2658.19
19	1205953166	423014	332365711	332141762	582741	2658.93	2657.13
20	1205953167	421941	331458110	331232089	607225	2651.66	2649.86
21	1205953168	422036	341350198	341127365	632384	2730.80	2729.02
22	1205953169	421595	342128682	341903486	657281	2737.03	2735.23
23	1205953170	420715	342053757	341820207	681738	2736.43	2734.56
24	1205953171	420220	345334964	345115990	706263	2762.68	2760.93
25	1205953172	420206	341921928	341704983	730791	2735.38	2733.64
26	1205953173	420475	342296971	342075095	755618	2738.38	2736.60
27	1205953174	420221	345722358	345502845	779978	2765.78	2764.02
28	1205953175	420712	345893615	345676653	804875	2767.15	2765.41
29	1205953176	421007	348752438	348530346	829483	2790.02	2788.24
30	1205953177	421165	340538589	340320060	854420	2724.31	2722.56
31	1205953178	421625	341623101	341402680	879740	2732.98	2731.22
32	1205953179	421577	346574694	346353471	904530	2772.60	2770.83
33	1205953180	421632	345105234	344884085	929167	2760.84	2759.07
34	1205953181	421917	341404203	341181189	954227	2731.23	2729.45
35	1205953182	422223	342547949	342318814	979454	2740.38	2738.55

Figura 3-3: Ejemplo salida del programa (fichero 2)

3.3.3 Estableciendo los Parámetros de Flujo

De acuerdo a lo comentado en apartados anteriores, los paquetes pertenecientes a un mismo flujo serán aquellos que compartan la misma quíntupla por la que son definidos, esto es, IP origen, IP destino, puerto origen, puerto destino y protocolo.

Además, como ya establecíamos en una de las entradas del programa, para que un paquete se clasifique dentro de un mismo flujo no sólo tiene que cumplir la quíntupla mencionada, sino que también deberá ajustarse al tiempo de expiración establecido.

El tiempo de expiración puede variar de acuerdo al propio fabricante del equipo Netflow. Esto se debe a la agregación de los datos que uno necesite o quiera aportar, por lo que encontraremos tiempos que oscilarán entre los 15 y 120 segundos.

Comenzando por Cisco, y política que hemos seguidos en este trabajo, ya sea por ser pionero en este tipo de tecnología e inspiración para este documento, se observa que establece como intervalo por defecto 15 segundos [9].

Otros, como el informe [18] utiliza los 30 segundos como tiempo de expiración, el mismo empleado por la red suiza SWITCH, sin embargo, sin una justificación clara del porqué de su decisión tomada.

Por otro lado, el informe [19] menciona un intervalo marcado en los 60 segundos, así como el informe [20], cercano al anterior estableciéndolo en los 64 segundos, al igual que el documento [21] cuando emplea un tiempo fijo.

Finalmente, algunos estudios mencionados anteriormente como el [12], utiliza un tiempo de 120 segundos, basándose en la referencia dada por el informe [22] y el informe [23].

3.3.4 Matlab

Una vez obtenidos los registros de red proporcionados por nuestro simulador en las trazas de CAIDA, requerimos del uso de otra herramienta que nos permitiera hacer el análisis de los datos reportados y extraer métricas (en nuestro caso, relación de ancho de banda y flujos activos), no solamente con el empleo directo de los informes, sino también la realización de cálculos a partir de éstos.

Para realizar el análisis de las estadísticas, disponíamos de múltiples herramientas disponibles en el mercado, sin embargo, reducimos las opciones a dos *softwares*: Matlab [24] y R [25], ambos no conocidos al inicio de este trabajo y decantándonos finalmente por el primero.

Como mencionábamos en secciones anteriores, las trazas proporcionadas por CAIDA se encontraban organizadas en enlaces, a su vez en años, dentro de éstos en días y, finalmente, en direcciones. Por ello, han sido múltiples los informes proporcionados por nuestro simulador y, por consiguiente, el análisis también ha sido un proceso costoso en cuanto al tiempo se refiere.

Hemos realizado diferentes *scripts* para realizar un análisis lo más completo posible de las estadísticas generadas, desde uno para el estudio de los informes diarios hasta otros para un estudio global de todos los datos.

Partiendo desde el estudio diario, ya que es el que nos reportará en un archivo de texto las características de cada uno de los informes generados. Por tanto, el archivo de texto contendrá los siguientes datos, siempre clasificados por el enlace, la fecha, la dirección y el protocolo (ejemplo en la Figura 3-4):

- Media de flujos.
- Media ancho de banda.
- Media del promedio de flujos a 5 minutos.
- Media del promedio ancho de banda a 5 minutos.
- Varianza de flujos.
- Varianza del ancho de banda.

- *Ratio de flujos, o simplemente ratio*. División de número de flujos activos entre el ancho de banda en Mb/s. Esta es la métrica de interés, en este trabajo, pues responde a la relación entre la carga de flujos y la carga en ancho de banda.

MFlows	MBW	meanFlows300	meanBW300	VFlows	VBW	VMFlows	VMBW	ratioFinal	date	enlace	dira	protocol
4.3747e+05	2.5925e+03	4.3669e+05	2.5949e+03	2.3848e+09	3.8946e+04	2.3848e+09	3.8946e+04	1.6874e+02	20080319	0	1	0
1.8487e+05	2.4476e+03	1.8431e+05	2.4513e+03	9.6064e+08	3.1022e+04	9.6064e+08	3.1022e+04	7.5532e+01	20080319	0	1	6
2.2903e+05	1.2842e+02	2.2884e+05	1.2733e+02	3.4876e+08	1.2000e+03	3.4876e+08	1.2000e+03	1.7834e+03	20080319	0	1	17
2.8272e+05	9.8987e+02	2.8287e+05	9.8917e+02	2.1035e+08	1.5341e+03	2.1035e+08	1.5341e+03	2.8561e+02	20080430	0	0	0
4.1407e+05	2.9739e+03	4.1456e+05	2.9845e+03	2.8869e+08	1.1572e+05	2.8869e+08	1.1572e+05	1.3924e+02	20080430	0	1	0
1.4060e+05	9.1551e+02	1.4071e+05	9.1498e+02	1.9400e+08	1.4439e+03	1.9400e+08	1.4439e+03	1.5358e+02	20080430	0	0	6
1.9556e+05	2.8327e+03	1.9584e+05	2.8430e+03	8.0652e+07	1.0603e+05	8.0652e+07	1.0603e+05	6.9038e+01	20080430	0	1	6
1.3589e+05	7.0485e+01	1.3592e+05	7.0409e+01	7.0081e+06	1.3285e+01	7.0081e+06	1.3285e+01	1.9279e+03	20080430	0	0	17
1.9722e+05	1.1614e+02	1.9740e+05	1.1654e+02	7.4726e+07	2.2890e+02	7.4726e+07	2.2890e+02	1.6981e+03	20080430	0	1	17
2.5436e+05	8.6364e+02	2.5408e+05	8.6208e+02	7.7519e+07	2.5581e+03	7.7519e+07	2.5581e+03	2.9452e+02	20080515	0	0	0
5.3175e+05	5.1632e+03	5.3141e+05	5.1580e+03	1.1572e+08	3.4519e+04	1.1572e+08	3.4519e+04	1.0299e+02	20080515	0	1	0
1.0752e+05	7.9888e+02	1.0728e+05	7.9730e+02	3.4882e+07	2.4319e+03	3.4882e+07	2.4319e+03	1.3458e+02	20080515	0	0	6
2.1536e+05	4.9487e+03	2.1517e+05	4.9439e+03	5.8172e+07	3.1900e+04	5.8172e+07	3.1900e+04	4.3518e+01	20080515	0	1	6
1.3805e+05	6.2634e+01	1.3795e+05	6.2642e+01	9.6864e+06	1.2310e+01	9.6864e+06	1.2310e+01	2.2041e+03	20080515	0	0	17
2.9229e+05	1.9813e+02	2.9208e+05	1.9794e+02	3.6295e+07	1.4843e+02	3.6295e+07	1.4843e+02	1.4752e+03	20080515	0	1	17
2.8371e+05	1.0624e+03	2.8378e+05	1.0628e+03	4.9257e+07	3.8702e+03	4.9257e+07	3.8702e+03	2.6704e+02	20080619	0	0	0
7.9882e+05	5.8084e+03	7.9853e+05	5.8120e+03	3.5006e+09	3.5062e+06	3.5006e+09	3.5062e+06	1.3753e+02	20080619	0	1	0
1.1559e+05	9.8963e+02	1.1548e+05	9.8969e+02	2.4693e+07	3.5574e+03	2.4693e+07	3.5574e+03	1.1681e+02	20080619	0	0	6
3.1914e+05	5.5369e+03	3.1884e+05	5.5405e+03	1.7943e+08	3.1873e+06	1.7943e+08	3.1873e+06	5.7639e+01	20080619	0	1	6
1.5984e+05	7.0715e+01	1.6001e+05	7.1029e+01	2.7617e+07	3.1838e+01	2.7617e+07	3.1838e+01	2.2603e+03	20080619	0	0	17
4.4422e+05	2.4414e+02	4.4424e+05	2.4422e+02	2.1070e+09	6.2173e+03	2.1070e+09	6.2173e+03	1.8195e+03	20080619	0	1	17
2.5459e+05	7.4882e+02	2.5310e+05	7.5106e+02	1.1995e+08	8.0435e+02	1.1995e+08	8.0435e+02	3.3999e+02	20080717	0	0	0
1.1592e+06	4.6487e+03	1.1588e+06	4.6478e+03	2.8755e+10	8.5450e+06	2.8755e+10	8.5450e+06	2.4936e+02	20080717	0	1	0
9.1213e+04	6.8119e+02	8.9801e+04	6.8349e+02	1.2225e+08	7.7190e+02	1.2225e+08	7.7190e+02	1.3390e+02	20080717	0	0	6
4.7659e+05	4.3714e+03	4.7583e+05	4.3712e+03	2.4271e+09	7.5574e+06	2.4271e+09	7.5574e+06	1.0902e+02	20080717	0	1	6
1.5520e+05	6.5877e+01	1.5515e+05	6.5812e+01	8.8013e+06	8.0983e+00	8.8013e+06	8.0983e+00	2.3559e+03	20080717	0	0	17
6.3343e+05	2.4886e+02	6.3368e+05	2.4846e+02	1.3474e+10	2.4517e+04	1.3474e+10	2.4517e+04	2.5453e+03	20080717	0	1	17
2.4749e+05	9.3847e+02	2.4653e+05	9.3463e+02	2.3794e+08	5.8614e+03	2.3794e+08	5.8614e+03	2.6372e+02	20080821	0	0	0
6.8557e+05	4.7046e+03	6.8529e+05	4.7020e+03	5.8000e+08	2.0741e+04	5.8000e+08	2.0741e+04	1.4572e+02	20080821	0	1	0
9.5148e+04	8.5729e+02	9.4708e+04	8.5347e+02	3.8056e+07	5.5454e+03	3.8056e+07	5.5454e+03	1.1099e+02	20080821	0	0	6
2.3345e+05	4.4187e+03	2.3311e+05	4.4160e+03	4.2284e+07	1.9575e+04	4.2284e+07	1.9575e+04	5.2832e+01	20080821	0	1	6
1.4504e+05	7.2754e+01	1.4452e+05	7.2722e+01	1.6625e+08	3.9246e+01	1.6625e+08	3.9246e+01	1.9935e+03	20080821	0	0	17
4.2727e+05	2.5269e+02	4.2738e+05	2.5281e+02	5.3314e+08	1.7236e+02	5.3314e+08	1.7236e+02	1.6909e+03	20080821	0	1	17
2.7490e+05	1.2093e+03	2.7449e+05	1.2082e+03	5.6545e+07	2.4738e+03	5.6545e+07	2.4738e+03	2.2731e+02	20080918	0	0	0
6.4823e+05	3.7013e+03	6.4844e+05	3.7028e+03	1.1254e+09	1.0397e+05	1.1254e+09	1.0397e+05	1.7514e+02	20080918	0	1	0
9.5781e+04	1.1114e+03	9.5460e+04	1.1101e+03	4.5078e+07	2.6890e+03	4.5078e+07	2.6890e+03	8.6180e+01	20080918	0	0	6
2.2593e+05	3.4524e+03	2.2605e+05	3.4549e+03	3.5428e+07	9.0901e+04	3.5428e+07	9.0901e+04	6.5442e+01	20080918	0	1	6
1.7236e+05	8.7066e+01	1.7232e+05	8.7257e+01	7.9854e+06	3.8260e+01	7.9854e+06	3.8260e+01	1.9796e+03	20080918	0	0	17
3.9766e+05	2.1657e+02	3.9775e+05	2.1564e+02	9.3688e+08	6.2388e+02	9.3688e+08	6.2388e+02	1.8362e+03	20080918	0	1	17

Figura 3-4: Ejemplo salida script Matlab

A partir de los datos proporcionados por los *scripts* implementados en Matlab, se realizará un análisis global, esto es, por enlace, dirección, año, etc. para comparar otros aspectos.

Además, a partir de cada uno de los *scripts* obtendremos también su correspondiente gráfica para poder evaluar también de modo visual el comportamiento de las diferentes estadísticas calculadas.

4 Análisis de resultados

4.1 Introducción

En esta sección realizaremos un estudio de las estadísticas y gráficas generadas previamente a través de los diferentes *scripts* desarrollados en Matlab y comentados en la sección anterior.

En primer lugar, realizaremos una explicación sobre el porqué es válido aplicar la media aritmética sobre las diferentes variables que fueron analizadas, con la ayuda del coeficiente de variación.

Posteriormente, veremos que en nuestro caso, realizar una distinción entre los flujos y el promedio de éstos a cinco minutos no nos arrojará ninguna diferencia significativa. Dicha distinción se pensó hacer en su momento con el fin de eliminar la *rafagosidad* propia del ancho de banda.

Además, esto último nos servirá para localizar las zonas que contienen un número de flujos normal, frente a aquellas que consideraremos anómalas y su correspondiente explicación.

Finalmente, realizaremos tanto un estudio espacial a lo largo de los años, como un estudio longitudinal de los diferentes enlaces disponibles y sus distintas direcciones que los componen.

4.2 Análisis de resultados

Para la discusión de los resultados nos hemos basado en las estadísticas obtenidas a través de la ejecución de los *scripts* de Matlab, los cuales no solo nos proporcionaban estadísticas, sino también gráficas con el fin de facilitar visualmente la observación de éstas. Un ejemplo de salida visual obtenida a través del *script* de Matlab la tenemos en la Figura 4-1.

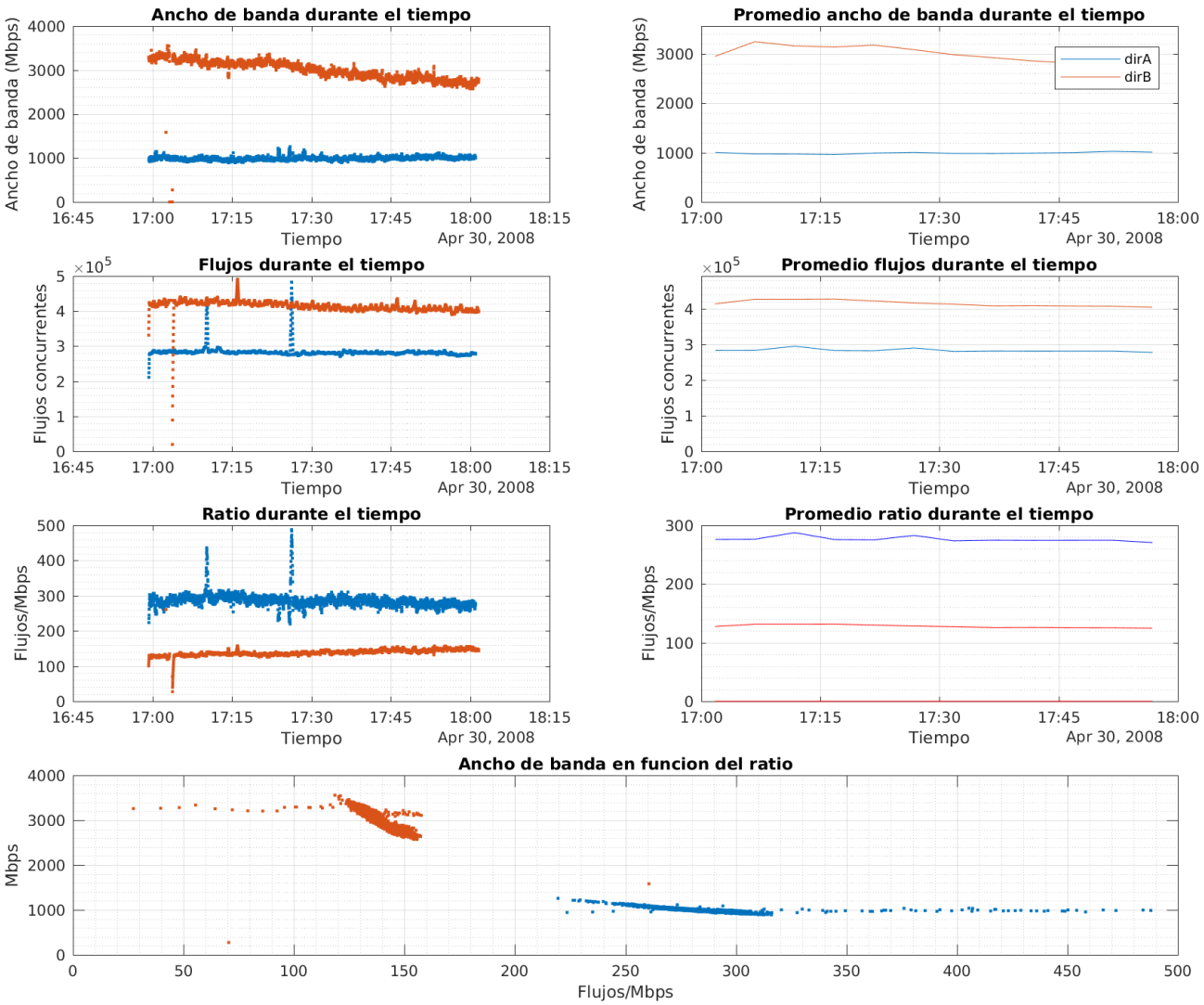


Figura 4-1: Ejemplo gráfica generada en Matlab para el 30 de abril de 2008 en el enlace de Chicago (ver Anexo B)

Las gráficas que mostraremos a continuación estarán basadas en las medias de los diferentes valores correspondientes a los flujos, ancho de banda, ratio (flujos dividido del ancho de banda), etc.

Con el fin de averiguar cuánto son de significativos los valores obtenidos tras aplicar la media a cada una de las variables, hemos calculado el coeficiente de variación tanto a los flujos de red como al ancho de banda.

A modo de recordatorio, la fórmula correspondiente al coeficiente de variación es la siguiente:

$$Cv = \frac{\sigma}{\bar{x}}$$

donde, σ es la desviación típica, esto es, la raíz cuadrada de la varianza, y $\bar{x}$ es la media.

El coeficiente de variación nos deberá arrojar valores menores que uno para poder considerar la media como una medida adecuada.

Por tanto, como podemos observar en la Figura 4-2, lo mencionado anteriormente se cumple tanto para los flujos como para el ancho de banda en el enlace de Chicago y, como veremos más adelante, las diferencias entre enlaces son marginales, por lo que también podremos considerarle esta medida.

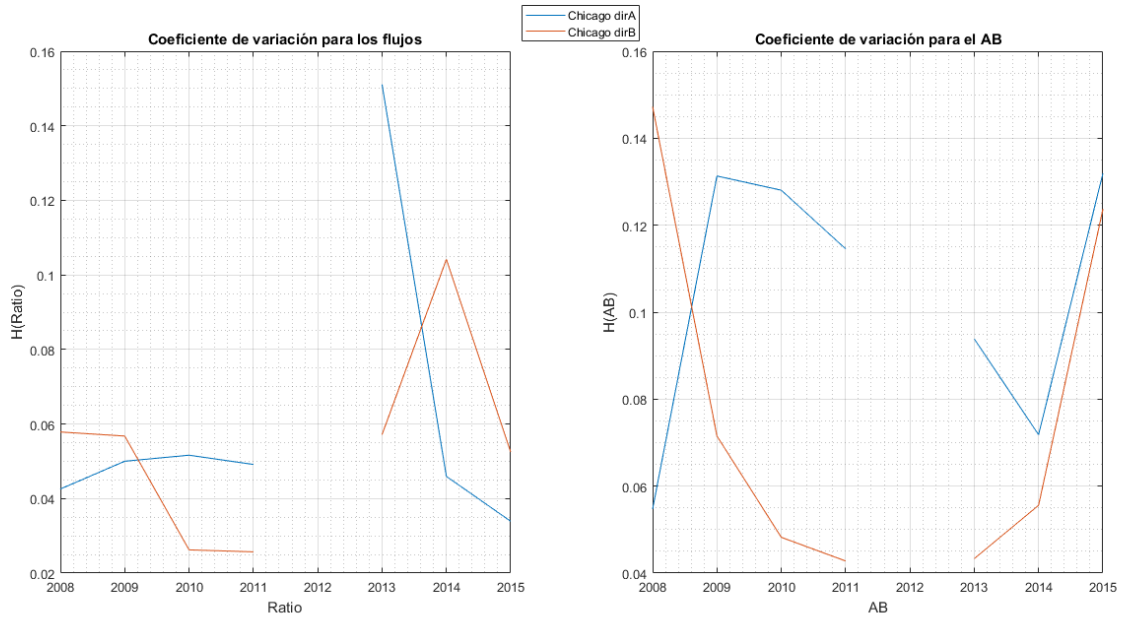


Figura 4-2: Coeficiente de variación para los flujos y ancho de banda en el enlace de Chicago

Como ya sabremos, o podemos imaginar, la agregación de los datos correspondientes al ancho de banda proporciona unos resultados que presentan, por lo general, una gran *rafagosidad* debido a su propia naturaleza. Por ello, quisimos tener en cuenta, no solo los valores proporcionados por nuestro programa, sino también realizar un promedio a cinco minutos (o 300 segundos) con el fin de reducir este fenómeno y obtener unos resultados más claros en cuanto a lo visual se refiere.

Sin embargo, y como vemos en la Figura 4-3 que muestra los valores de ratio de flujos para todas las trazas en análisis arroja valores indistinguibles respecto al promedio

de cinco minutos y un segundo, lo que nos lleva a pensar que, en nuestro caso, no tendremos problemas respecto a lo mencionado anteriormente.

Más relevante de esta Figura 4-3 es que nos va a servir para localizar las zonas de influencia normal de la variable ratio de flujos objeto de este estudio, así como terminar las muestras que consideraremos fuera de esta normalidad las cuales merecen un estudio más detallado si cabe.

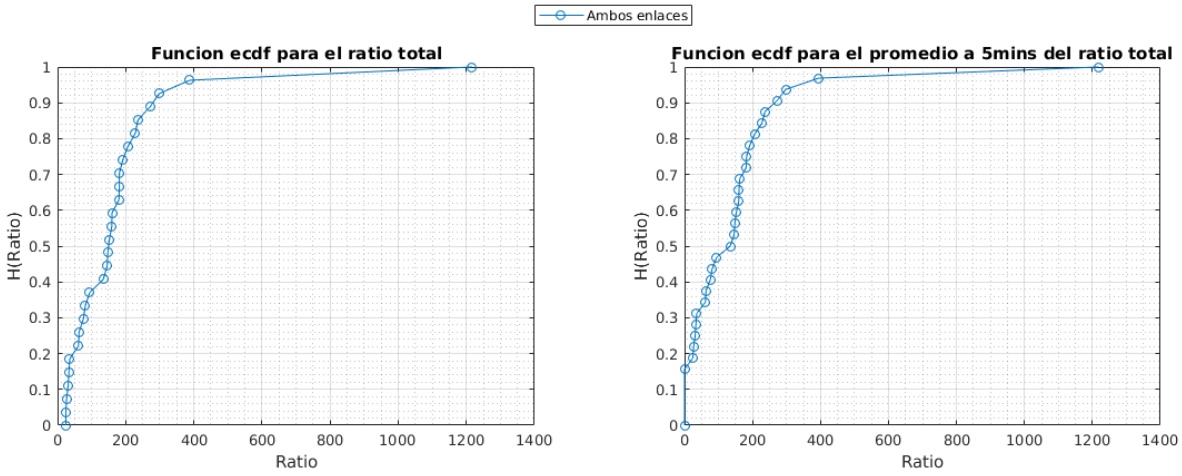


Figura 4-3: Comparación función ECDF para las medias del ratio de flujosen ambos enlaces en ambos sentidos

El primer resultado de interés es localizar más del 90% de los puntos entre los ratios (recuerde, número de flujos divido entre el ancho de banda) 0 y 300. De este modo, ante un sistema basado en flujo conectado a un enlace comercial y en hora cargada (esto es, las trazas de CAIDA en estudio) podemos estimar que cada Mb/s transporta típicamente menos de 300 flujos y más de la mitad menos de 150 flujos. Estas cifras son las que un gestor de red puede esperar en escenario equivalentes.

Por el contrario, de forma anómala, se encuentran muestras en las que un solo Mb/s transporta más de 1200 flujos. Al ver este punto tan desplazado del resto, lo consideramos como una anomalía y procedimos a investigar sobre ello.

De modo que, realizamos una distinción entre enlaces y, a su vez, entre direcciones para situar este punto de una manera más exacta, tal y como vemos en la Figura 4-4.

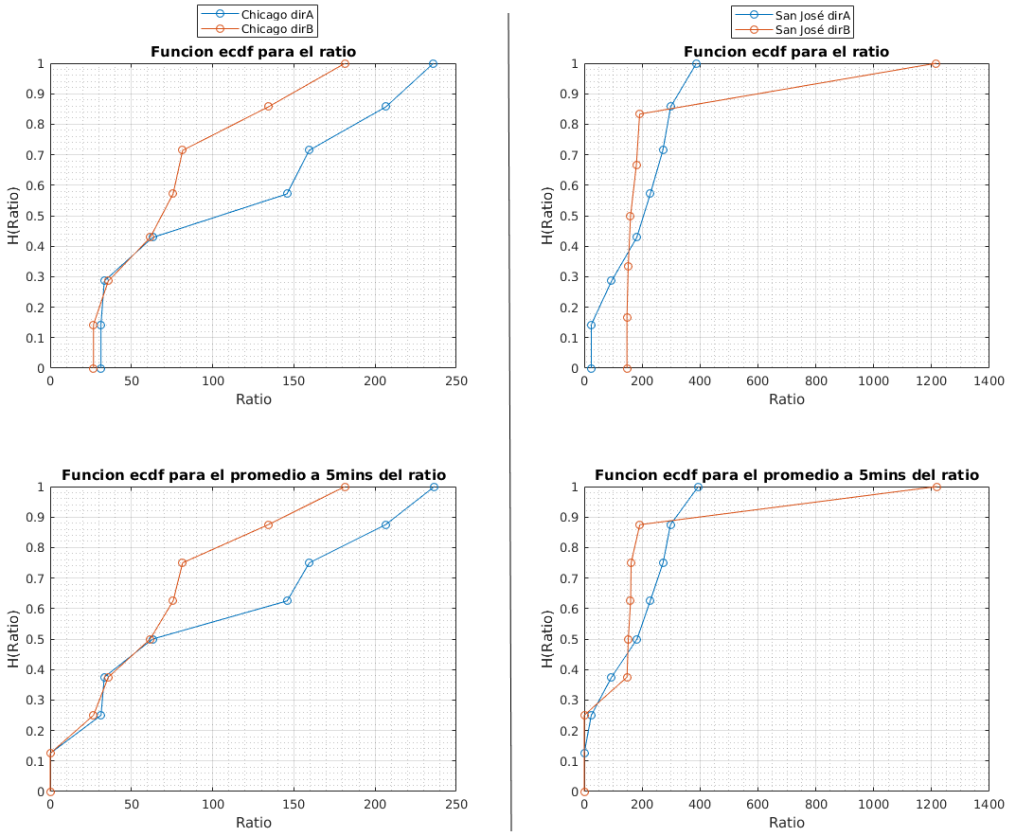


Figura 4-4: Comparación de función ECDF entre enlaces y direcciones

Una vez ubicado el punto anómalo, enlace de San José en la dirección procedente de Los Ángeles hacia San José, barajamos dos hipótesis: un aumento de los flujos de red o un aumento del ancho de banda y, al comprobar la salida del *script* de Matlab, observamos que para el día 17 de octubre de 2011 se produce un incremento en los flujos concurrentes con respecto al resto de días de los diferentes años disponibles.

Este incremento puede ser ocasionado por diferentes elementos: las aplicaciones que se están ejecutando en el día indicado, el tipo de usuarios que están haciendo uso de la red o, al tratarse de una red comercial, un cambio en la proporción de usuarios "domésticos" o profesionales [26]. Y de forma menos optimista a algún tipo de ataque o fallo en la red [1].

Por tanto, podemos de nuevo reportar que en general el ratio habitual se localizará entre el 0 y los 250 y, de una forma más generosa, entre el 0 y los 400. Además, calificaremos el punto distante del resto como una zona anormal debido a la diferencia que se produjo en el número de flujos. De hecho, este estudio puede ser de interés para que el gestor de red preste atención a lo sucedido ese día.

A continuación, realizaremos tanto un estudio longitudinal del ratio (flujos dividido entre ancho de banda), del ancho de banda y los flujos de red con el fin de observar el desarrollo que han tenido las diferentes variables a lo largo de los años. También consideraremos un estudio espacial a modo de comparativa entre los diferentes enlaces disponibles, y mencionados en secciones anteriores, para observar sus posibles diferencias.

Las gráficas que se encuentras al lado izquierdo de la línea horizontal de la Figura 4-5, se corresponde con el enlace localizado en Chicago, Illinois. En ésta, podremos observar las dos direcciones de ida y vuelta en las que se divide el enlace, siendo la dirección marcada bajo la letra *A* la que inicia en Chicago y parte hacia Seattle, Washington, y la dirección *B* la que inicia en Seattle y termina en Chicago. Su correspondiente gráfica promedio a cinco minutos es la que se encuentra justo debajo de ella misma.

Por otro lado, las gráficas que se encuentras al lado derecho de la línea horizontal de la Figura 4-5, se corresponde con el enlace localizado en San José, California. En ésta podremos observar las dos direcciones de ida y vuelta en las que se divide el enlace, siendo la dirección marcada bajo la letra *A* la que inicia en San José y parte hacia Los Ángeles, California, y la dirección *B* la que inicia en Los Ángeles y termina en San José. Su correspondiente gráfica promedio a cinco minutos es la que se encuentra justo debajo de ella misma.

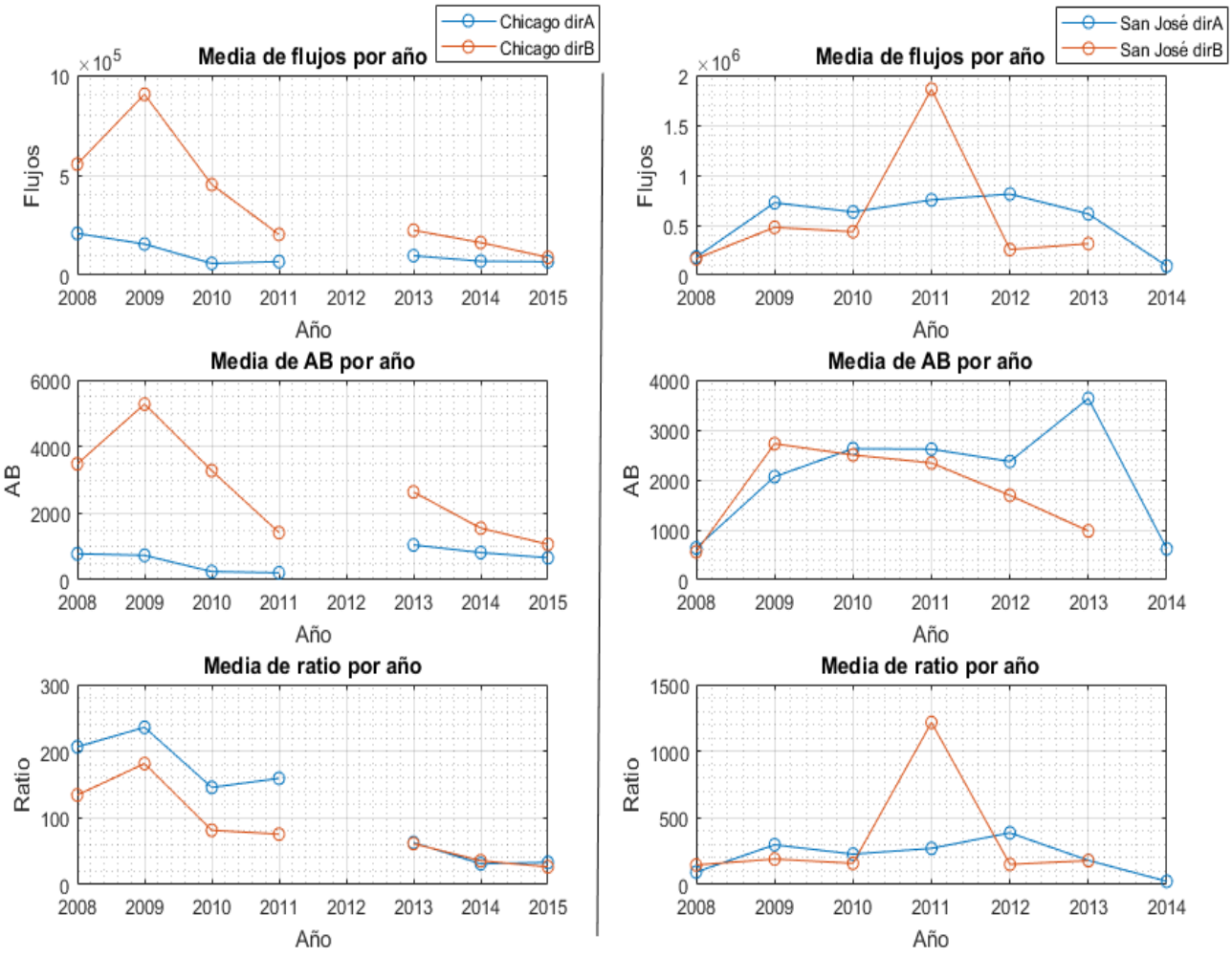


Figura 4-5: Estudio longitudinal y espacial

El espacio en blanco que podremos ver en el año 2012 para el enlace de Chicago y, en el año 2014 para el enlace de San José en la dirección *B*, se debe a la falta de datos en las

trazas proporcionadas por CAIDA, tal y como se comentó en la parte de disponibilidad de datos.

Una vez más, veremos reflejado en la Figura 4-5 el resultado del incremento de flujos que se comentó anteriormente, pero obviando esto, podremos comentar las tendencias de cada uno de los enlaces.

En el enlace de Chicago se observa una clara tendencia decreciente a lo largo de los años, tanto en el número de flujos de red como en el ancho de banda consumido, mientras que en el enlace de San José obtenemos una querencia plana en ambos sentidos. Por otro lado, en San José ambos sentidos se comportan de manera similar, mientras que en Chicago su comportamiento es en menor medida. En definitiva cierta heterogeneidad. Cuando comparamos el ratio de flujos entre ambos enlaces, ignorando huecos y valores anómalos, vemos un comportamiento notablemente homogéneo como ya vimos en figuras anteriores.

Idealmente estos cambios, más allá de la anomalía detectada, creemos que pueden deberse, como se comentaba con anterioridad al cambio del mix de aplicaciones que transporta el enlace. Para ello sería de interés clasificar el tráfico, sin embargo las trazas de CAIDA no contienen carga útil, dificultando cualquier clasificación fidedigna, por ejemplo, basada en DPI [2].

5 Conclusiones

5.1 Introducción

En esta sección se plasmarán las conclusiones finales a tomar tras la realización de este Trabajo de Fin de Grado.

5.2 Conclusiones

Como ya sabremos, el comportamiento de las redes viene dado por patrones que se ajustan a las características de éstas mismas, esto es, las aplicaciones que se ejecutan sobre la red, el tipo de usuarios que hacen uso de ella: usuarios de empresa o domésticos, o el tipo de red de la que realizaremos el análisis: doméstica o profesional.

Por todo ello, y unido al crecimiento exponencial del uso de la Internet en los últimos años, la monitorización y análisis de la red ha ido creciendo en importancia debido a información que es capaz de proporcionarnos, especialmente para aquellas empresas proveedoras de servicios en la Internet (ISP), capaces de abordar de una forma más rauda y exacta aspectos relacionados con QoS o vulnerabilidades, como DoS.

Debido al aumento de las velocidades de transmisión en las redes actuales, la monitorización y análisis de paquetes se ha convertido en una opción muy difícil a día de hoy por diferentes motivos, adquiriendo importancia los flujos de red en los sistemas encargados de ello.

Un flujo de red consiste en un conjunto de paquetes que pasan por un punto durante un intervalo de tiempo y que comparten cinco atributos: dirección origen y destino IP, puerto origen y destino, y la capa de transporte. Éstos nos proporcionarán la mayor información posible acerca de un conjunto de paquetes.

Para la obtención de toda esta información se emplean sistemas de monitorización y análisis de tráfico en tiempo real cuyas arquitecturas están conformadas por tres componentes de procesamiento: generador de flujos, contenedor de flujos y el analizador del tráfico. Un ejemplo de éstos es NetFlow, creado por Darren y Brarry Bruins, empleados de Cisco, en el año 1996.

Para este Trabajo de Fin de Grado se han obtenido todas las trazas del portal oficial del Centro para el Análisis de Datos Aplicados de la Internet (CAIDA), cuyos objetivos principales son la investigación de la red y la construcción de infraestructuras de investigación.

Todas estas trazas están divididas en dos enlaces diferentes, uno de ellos localizado en Chicago, Illinois, y el otro en San José, California, ambos ubicados en Estados Unidos, englobadas en varias decenas de Terabytes que fueron tratadas posteriormente por un programa de análisis.

Este programa ha sido desarrollado por nosotros mismos y se ha cuidado en todo momento tanto el empleo de los recursos de los que hacía uso como los tiempos de ejecución, llegando a procesar cerca de 3000 millones de paquetes por segundo. Éste nos proporcionará un informe con los diferentes datos analizados como son: flujos concurrentes, ancho de banda consumido, tiempo de llegada de los paquetes, etc.

Dichos informes proporcionados por el programa son utilizados posteriormente por nuestros *scripts* ejecutados en Matlab, los cuales nos arrojarán diferentes estadísticas acerca de los valores obtenidos y diferentes gráficas capaces de facilitarnos visualmente la comprensión de los resultados.

Entre todas estas gráficas queremos destacar una de ellas, la Figura 4-3, la cual nos servirá para aportar finalmente la conclusión que nos responderá a la pregunta que aborda nuestro Trabajo de Fin de Grado, que es la siguiente: ¿dada una red comercial cuántos flujos concurrentes podremos esperar para un ancho de banda consumido?

Como observamos, el 85% de las medidas, que hemos denominado ratio, están en el intervalo de 0 a 250, el cual consideraremos como zona normal de operación. Rango que se puede ampliar a 400, para alcanzar más del 95% de las muestras. Esto resume que puede esperarse en un escenario habitual.

Por otro lado, hemos encontrado un punto anómalo habiéndose ocasionado por un crecimiento anormal de flujos concurrentes, originado por diferentes motivos: desde un crecimiento de usuarios hasta la explotación de una vulnerabilidad de nuestra red. Esto debe traducirse como un escenario anómalo, siendo decisión del gestor de red dimensionar la red para un escenario habitual con control de anomalías, o ir más allá, y sobredimensionar el equipo para que aun en escenario anómalo los sistemas funcionen sin problema.

Finalmente hemos encontrado notable heterogeneidad cuando hemos considerado factores como el sentido del tráfico y el tiempo. Por un lado, un enlace ha mostrado un decaimiento en el ratio de flujos a lo largo del tiempo mientras otro enlace se ha mostrado más insensible. Por otro lado, hemos visto que en el caso de San José los sentidos son muy similares entre ellos pero no así en los de Chicago.

6 Trabajo futuro

6.1 Introducción

En esta sección se proporcionará una lista de posibles trabajos futuros a realizar para ampliar el análisis de los flujos de red.

6.2 Trabajo futuro

Para abarcar funcionalidad que no se presenta en este trabajo y con ello extender el estudio y análisis de los flujos de red, dejamos a continuación una lista de posibles trabajos a realizar en un futuro. Entre ellos se encuentran:

- Extender a más redes. En este Trabajo de Fin de Grado se han empleado las trazas proporcionadas por CAIDA, que se corresponden con los enlaces localizados en Chicago, Illinois, y San José, California, ambos ubicados en Estados Unidos.

  Por tanto, extender el trabajo a más y diferentes redes nos proporcionará no solo el conocimiento del comportamiento de la red, sino también la posibilidad de comparar redes localizadas en mismos o diferentes países, entre redes domésticas o profesionales, los diferentes usos posibles de las redes, etc.

- Extender a diferentes tiempos de expiración. En este Trabajo de Fin de Grado se fijó el tiempo de expiración de los flujos concurrentes de red en 15 segundos, pero como ya comentamos, existen otros valores posibles.

  Ampliar el tiempo de expiración nos permitirá obtener una agregación de datos mayor en las diferentes variables a estudio y obtener con ello una visión ligeramente distinta de los datos originales.

- Como ya comentamos anteriormente, los paquetes de las trazas proporcionadas por CAIDA no contienen la carga útil con el fin de anonimizarlas, limitando así el uso de técnicas DPI. Por ello, no se realizó un estudio distinguiendo las diferentes aplicaciones posibles que se encontrarían detrás de cada enlace, además de que no se correspondería con el tema principal de este Trabajo de Fin de Grado.

  Por ello, se propone utilizar otras técnicas basadas en estadísticas que aunque menos precisas nos den una intuición de que aplicaciones están generando más o menos flujos.

Referencias

[1] Victor Moreno, Pedro M. Santiago del Río, Javier Ramos, David Muelas, José Luis García-Dorado, Francisco J. Gómez-Arribas, Javier Aracil. Multi-granular, multi-purpose and multi-Gb/s monitoring on off-the-shelf systems. International Journal of Network Management. 2014.

[2] Thuy T.T. Nguyen, Grenville Armitage. A Survey of Techniques for Internet Traffic Classification using Machine Learning. IEEE Communications Surveys & Tutorials. 2008.

[3] José Luis García-Dorado, Pedro M. Santiago del Río, Javier Ramos, David Muelas, Victor Moreno, Jorge E. López de Vergara, Javier Aracil. Low-cost and High-performance: VoIP monitoring and full-data retention at multi-Gb/s rates using commodity hardware. International Journal of Network Management. 2014.

[4] Sally Floyd, Vern Paxson. Difficulties in Simulating the Internet. IEEE/ACM Transactions on Networking. 2001.

[5] Zhang Weiwei, Gong Jian, Gu Wenjie, Cai Shaomin. NetFlow-Based Network Traffic Monitoring. Asia-Pacific Network Operations and Management Symposium. 2011.

[6] N. Brownlee. Traffic Flow Measurement: Experience with NeTraMet, IETF RFC2123. 2007.

[7] Luca Deri. Ntop. Disponible en http://www.ntop.org/

[8] Se-Hee Han, Myung-Sup Kim, Hong-Taek Ju, J.W. Hong. The Architecture of NG-MON: A Passive Network Monitoring System for High-Speed IP Networks. Distributed Systems: Operations and Management Conference. 2002.

[9] CISCO. Introduction to Cisco IOS NetFlow. 2012.

[10] Internet Protocol Flow Information eXport.
https://datatracker.ietf.org/wg/ipfix/charter/

[11] Bingdong Li, Jeff Springer, George Bebis, Mehmet Hadi Gunes. A survey of network flow applications.. Journal of Network and Computer Applications. 2013.

[12] Myung-Sup Kim, Young J. Won, James W. Hong. Characteristic analysis of Internet traffic from the perspective of flows. Computer Communications. 2006.

[13] Localización de las plataformas de medición de CAIDA. Disponible en http://www.caida.org/projects/ark/locations/

[14] Página oficial de CAIDA. Disponible en http://www.caida.org/

[15] Página oficial de Equinix. Disponible en http://www.equinix.com/

[16] J.L. Garcia-Dorado, J.A. Hernández, J. Aracil, J.E. Lopez de Vergara, y Sergio Lopez-Buedo. Characterization of the busy-hour traffic of IP networks based on their intrinsic features. Computer Networks. 2011.

[17] John Kristoff. Northwestern University. Disponible en http://www.cc.gatech.edu/computing/Networking/projects/cryptopan/

[18] Daniela Brauckhoff, Bernhard Tellenbach, Arno Wagner, Martin May, Anukool Lakhina. Impact of Packet Sampling on Anomaly Detection Metrics. ACM SIGCOMM Conference on Internet Measurement. 2006.

[19] Cristian Estan, Ken Keys, David Moore, George Varghese. Building a Better Netflow. Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. 2004.

[20] W.J. Lui, J.Gong. Double Sampling for Flow Measurement on High Speed Links. Computer Networks. 2008.

[21] Nicolas Hohn, Darryl Veitch. Inverting Sampled Traffic. IEEE/ACM Transactions on Networking. 2006.

[22] Gianlucca Iannaccone, Christophe Diot, Ian Graham, Nick McKeown. Monitoring very High Speed Links. ACM SIGCOMM Workshop on Internet measurement. 2001.

[23] Andrea Bianco, Gianluca Mardence, Marco Mellia, Maurizio M. Munafò, Luca Muscariello. Web User-Session Interference by Means of Clustering Techniques. 2005.

[24] Página oficial de Matlab. Disponible en https://es.mathworks.com/

[25] Página oficial de R. Disponible en https://www.r-project.org

[26] F. Mata, P. Żuraniewski, M. Mandjes, M. Mellia. Anomaly detection in diurnal data. Computer Networks. 2014.

Glosario

ISP *Internet Service Provider* (Proveedor de Servicios de Internet). Empresa que ofrece conexión a Internet a sus clientes.

QoS *Quality of Service* (Calidad de Servicio). Rendimiento promedio de una red de telefonía o de computadoras, particularmente el rendimiento visto por los usuarios de la red.

DoS *Denial of Service* (Denegación del Servicio). Ataque basado en efectuar peticiones muy continuadas a un servidor con el propósito de dejarlo inutilizado.

IP *Internet Protocol* (Protocolo de Internet). Protocolo de comunicación clasificado en la capa de red según el modelo OSI.

CAIDA *Center for Applied Internet Data Analysis* (Centro de Análisis de Datos Aplicados de Internet). Es una institución colaboradora con organizaciones de los sectores comercial, gubernamental e investigador, con el objetivo de promover una mayor cooperación en la ingeniería y el mantenimiento de una infraestructura global robusta y escalable de Internet.

TCP *Transmission Control Protocol* (Protocolo de Control de la Transmisión). Protocolo de transporte con garantías de recepción.

UDP *User Datagram Protocol* (Protocolo de Datagramas de Usuario). Protocolo del nivel de transporte basado en el intercambio de datagramas.

P2P *Peer-to-peer* (Red entre iguales). Es una red de ordenadores en la que todos o algunos aspectos funcionan sin clientes ni servidores fijos, sino una serie de nodos que se comportan como iguales entre sí.

WAN *Wide Area Network* (Red de área amplia). Es una red de computadoras que une varias redes locales, aunque sus miembros no estén todos en una misma ubicación física.

SNMP *Simple Network Management Protocol* (Protocolo Simple de Administración de Red). Es un protocolo de la capa de aplicación que facilita el intercambio de información de administración entre dispositivos de red.

Anexos

A. Tabla de archivos disponibles

Año	Enlace	Mes	Dirección	Disponibilidad
2008	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	-
			Seattle-Chicago	-
		03	Chicago-Seattle	-
			Seattle-Chicago	X
		04	Chicago-Seattle	X
			Seattle-Chicago	X
		05	Chicago-Seattle	X
			Seattle-Chicago	X
		06	Chicago-Seattle	X
			Seattle-Chicago	X
		07	Chicago-Seattle	X
			Seattle-Chicago	X
		08	Chicago-Seattle	X
			Seattle-Chicago	X
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	X
			Seattle-Chicago	X
		11	Chicago-Seattle	X
			Seattle-Chicago	X
		12	Chicago-Seattle	X
			Seattle-Chicago	X
	San José	01	San José-Los Ángeles	-
			Los Ángeles-San José	-
		02	San José-Los Ángeles	-
			Los Ángeles-San José	-
		03	San José-Los Ángeles	-

			Los Ángeles-San José	-
		04	San José-Los Ángeles	-
			Los Ángeles-San José	-
		05	San José-Los Ángeles	-
			Los Ángeles-San José	-
		06	San José-Los Ángeles	-
			Los Ángeles-San José	-
		07	San José-Los Ángeles	X
			Los Ángeles-San José	X
		08	San José-Los Ángeles	X
			Los Ángeles-San José	X
		09	San José-Los Ángeles	-
			Los Ángeles-San José	X
		10	San José-Los Ángeles	X
			Los Ángeles-San José	X
		11	San José-Los Ángeles	X
			Los Ángeles-San José	X
		12	San José-Los Ángeles	-
			Los Ángeles-San José	X

2009	Chicago	01	Chicago-Seattle	X
			Seattle-Chicago	X
		02	Chicago-Seattle	X
			Seattle-Chicago	X
		03	Chicago-Seattle	X
			Seattle-Chicago	X
		04	Chicago-Seattle	X
			Seattle-Chicago	X
		05	Chicago-Seattle	X
			Seattle-Chicago	X
		06	Chicago-Seattle	X
			Seattle-Chicago	X
		07	Chicago-Seattle	X
			Seattle-Chicago	X
		08	Chicago-Seattle	X

			Seattle-Chicago	X
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	X
			Seattle-Chicago	X
		11	Chicago-Seattle	X
			Seattle-Chicago	X
		12	Chicago-Seattle	X
			Seattle-Chicago	X
	San José	01	San José-Los Ángeles	-
			Los Ángeles-San José	X
		02	San José-Los Ángeles	-
			Los Ángeles-San José	X
		03	San José-Los Ángeles	X
			Los Ángeles-San José	X
		04	San José-Los Ángeles	X
			Los Ángeles-San José	X
		05	San José-Los Ángeles	X
			Los Ángeles-San José	X
		06	San José-Los Ángeles	X
			Los Ángeles-San José	X
		07	San José-Los Ángeles	X
			Los Ángeles-San José	X
		08	San José-Los Ángeles	X
			Los Ángeles-San José	X
		09	San José-Los Ángeles	X
			Los Ángeles-San José	X
		10	San José-Los Ángeles	X
			Los Ángeles-San José	X
		11	San José-Los Ángeles	X
			Los Ángeles-San José	X
		12	San José-Los Ángeles	X
			Los Ángeles-San José	X

2010	Chicago	01	Chicago-Seattle	X
			Seattle-Chicago	X
		02	Chicago-Seattle	X
			Seattle-Chicago	X
		03	Chicago-Seattle	X
			Seattle-Chicago	X
		04	Chicago-Seattle	X
			Seattle-Chicago	X
		05	Chicago-Seattle	-
			Seattle-Chicago	-
		06	Chicago-Seattle	-
			Seattle-Chicago	-
		07	Chicago-Seattle	-
			Seattle-Chicago	-
		08	Chicago-Seattle	X
			Seattle-Chicago	X
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	X
			Seattle-Chicago	X
		11	Chicago-Seattle	-
			Seattle-Chicago	-
		12	Chicago-Seattle	-
			Seattle-Chicago	-
	San José	01	San José-Los Ángeles	X
			Los Ángeles-San José	X
		02	San José-Los Ángeles	X
			Los Ángeles-San José	X
		03	San José-Los Ángeles	X
			Los Ángeles-San José	X
		04	San José-Los Ángeles	X
			Los Ángeles-San José	X
		05	San José-Los Ángeles	-
			Los Ángeles-San José	-
		06	San José-Los Ángeles	-

			Los Ángeles-San José	X
		07	San José-Los Ángeles	X
			Los Ángeles-San José	X
		08	San José-Los Ángeles	X
			Los Ángeles-San José	X
		09	San José-Los Ángeles	X
			Los Ángeles-San José	X
		10	San José-Los Ángeles	X
			Los Ángeles-San José	X
		11	San José-Los Ángeles	X
			Los Ángeles-San José	X
		12	San José-Los Ángeles	X
			Los Ángeles-San José	X

2011	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	X
			Seattle-Chicago	X
		03	Chicago-Seattle	X
			Seattle-Chicago	X
		04	Chicago-Seattle	X
			Seattle-Chicago	X
		05	Chicago-Seattle	X
			Seattle-Chicago	X
		06	Chicago-Seattle	-
			Seattle-Chicago	-
		07	Chicago-Seattle	X
			Seattle-Chicago	X
		08	Chicago-Seattle	X
			Seattle-Chicago	X
		09	Chicago-Seattle	-
			Seattle-Chicago	-
		10	Chicago-Seattle	-
			Seattle-Chicago	-
		11	Chicago-Seattle	-

			Seattle-Chicago	-
		12	Chicago-Seattle	-
			Seattle-Chicago	-
	San José	01	San José-Los Ángeles	X
			Los Ángeles-San José	X
		02	San José-Los Ángeles	X
			Los Ángeles-San José	X
		03	San José-Los Ángeles	X
			Los Ángeles-San José	X
		04	San José-Los Ángeles	X
			Los Ángeles-San José	X
		05	San José-Los Ángeles	X
			Los Ángeles-San José	X
		06	San José-Los Ángeles	-
			Los Ángeles-San José	-
		07	San José-Los Ángeles	X
			Los Ángeles-San José	X
		08	San José-Los Ángeles	X
			Los Ángeles-San José	X
		09	San José-Los Ángeles	X
			Los Ángeles-San José	X
		10	San José-Los Ángeles	X
			Los Ángeles-San José	X
		11	San José-Los Ángeles	X
			Los Ángeles-San José	X
		12	San José-Los Ángeles	X
			Los Ángeles-San José	X

2012	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	-
			Seattle-Chicago	-
		03	Chicago-Seattle	-
			Seattle-Chicago	-
		04	Chicago-Seattle	-

			Seattle-Chicago	-
		05	Chicago-Seattle	-
			Seattle-Chicago	-
		06	Chicago-Seattle	-
			Seattle-Chicago	-
		07	Chicago-Seattle	-
			Seattle-Chicago	-
		08	Chicago-Seattle	-
			Seattle-Chicago	-
		09	Chicago-Seattle	-
			Seattle-Chicago	-
		10	Chicago-Seattle	-
			Seattle-Chicago	-
		11	Chicago-Seattle	-
			Seattle-Chicago	-
		12	Chicago-Seattle	-
			Seattle-Chicago	-
	San José	01	San José-Los Ángeles	X
			Los Ángeles-San José	X
		02	San José-Los Ángeles	X
			Los Ángeles-San José	X
		03	San José-Los Ángeles	X
			Los Ángeles-San José	X
		04	San José-Los Ángeles	-
			Los Ángeles-San José	-
		05	San José-Los Ángeles	X
			Los Ángeles-San José	X
		06	San José-Los Ángeles	X
			Los Ángeles-San José	X
		07	San José-Los Ángeles	X
			Los Ángeles-San José	X
		08	San José-Los Ángeles	X
			Los Ángeles-San José	X
		09	San José-Los Ángeles	X
			Los Ángeles-San José	X

		10	San José-Los Ángeles	X
			Los Ángeles-San José	X
		11	San José-Los Ángeles	X
			Los Ángeles-San José	X
		12	San José-Los Ángeles	X
			Los Ángeles-San José	X

2013	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	-
			Seattle-Chicago	-
		03	Chicago-Seattle	-
			Seattle-Chicago	-
		04	Chicago-Seattle	-
			Seattle-Chicago	-
		05	Chicago-Seattle	X
			Seattle-Chicago	X
		06	Chicago-Seattle	X
			Seattle-Chicago	X
		07	Chicago-Seattle	X
			Seattle-Chicago	X
		08	Chicago-Seattle	X
			Seattle-Chicago	X
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	X
			Seattle-Chicago	X
		11	Chicago-Seattle	X
			Seattle-Chicago	X
		12	Chicago-Seattle	X
			Seattle-Chicago	X
	San José	01	San José-Los Ángeles	X
			Los Ángeles-San José	X
		02	San José-Los Ángeles	X
			Los Ángeles-San José	X

		03	San José-Los Ángeles	X
			Los Ángeles-San José	X
		04	San José-Los Ángeles	X
			Los Ángeles-San José	X
		05	San José-Los Ángeles	X
			Los Ángeles-San José	X
		06	San José-Los Ángeles	X
			Los Ángeles-San José	X
		07	San José-Los Ángeles	X
			Los Ángeles-San José	-
		08	San José-Los Ángeles	X
			Los Ángeles-San José	-
		09	San José-Los Ángeles	X
			Los Ángeles-San José	-
		10	San José-Los Ángeles	X
			Los Ángeles-San José	-
		11	San José-Los Ángeles	X
			Los Ángeles-San José	-
		12	San José-Los Ángeles	X
			Los Ángeles-San José	-

2014	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	-
			Seattle-Chicago	-
		03	Chicago-Seattle	X
			Seattle-Chicago	X
		04	Chicago-Seattle	-
			Seattle-Chicago	-
		05	Chicago-Seattle	-
			Seattle-Chicago	-
		06	Chicago-Seattle	X
			Seattle-Chicago	X
		07	Chicago-Seattle	-
			Seattle-Chicago	-

		08	Chicago-Seattle	-
			Seattle-Chicago	-
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	-
			Seattle-Chicago	-
		11	Chicago-Seattle	-
			Seattle-Chicago	-
		12	Chicago-Seattle	X
			Seattle-Chicago	X
	San José	01	San José-Los Ángeles	-
			Los Ángeles-San José	-
		02	San José-Los Ángeles	-
			Los Ángeles-San José	-
		03	San José-Los Ángeles	X
			Los Ángeles-San José	-
		04	San José-Los Ángeles	-
			Los Ángeles-San José	-
		05	San José-Los Ángeles	-
			Los Ángeles-San José	-
		06	San José-Los Ángeles	X
			Los Ángeles-San José	-
		07	San José-Los Ángeles	-
			Los Ángeles-San José	-
		08	San José-Los Ángeles	-
			Los Ángeles-San José	-
		09	San José-Los Ángeles	-
			Los Ángeles-San José	-
		10	San José-Los Ángeles	-
			Los Ángeles-San José	-
		11	San José-Los Ángeles	-
			Los Ángeles-San José	-
		12	San José-Los Ángeles	-
			Los Ángeles-San José	-

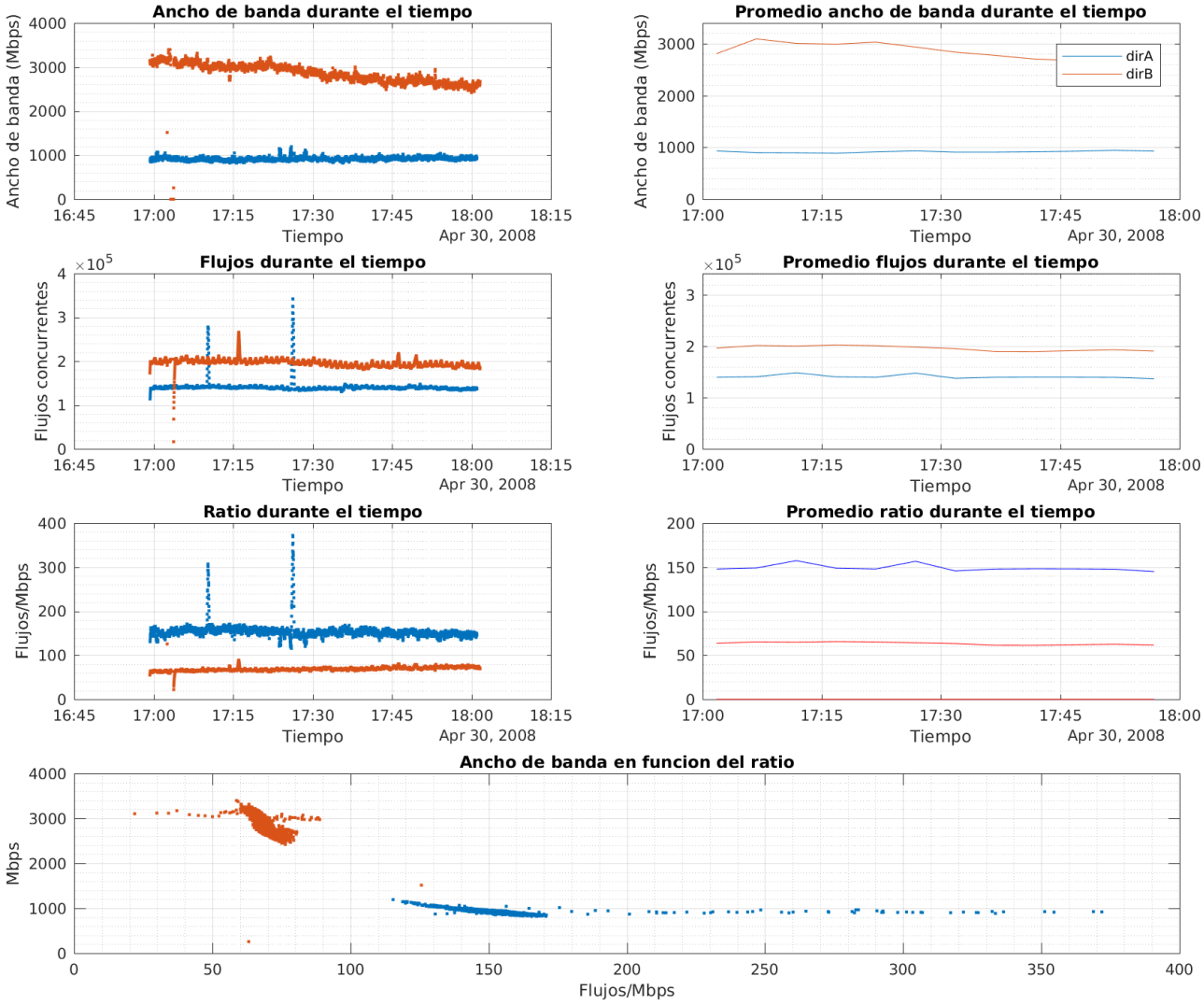
X

2015	Chicago	01	Chicago-Seattle	-
			Seattle-Chicago	-
		02	Chicago-Seattle	X
			Seattle-Chicago	X
		03	Chicago-Seattle	-
			Seattle-Chicago	-
		04	Chicago-Seattle	-
			Seattle-Chicago	-
		05	Chicago-Seattle	X
			Seattle-Chicago	X
		06	Chicago-Seattle	-
			Seattle-Chicago	-
		07	Chicago-Seattle	-
			Seattle-Chicago	-
		08	Chicago-Seattle	-
			Seattle-Chicago	-
		09	Chicago-Seattle	X
			Seattle-Chicago	X
		10	Chicago-Seattle	-
			Seattle-Chicago	-
		11	Chicago-Seattle	-
			Seattle-Chicago	-
		12	Chicago-Seattle	X
			Seattle-Chicago	X
	San José	01	San José-Los Ángeles	-
			Los Ángeles-San José	-
		02	San José-Los Ángeles	-
			Los Ángeles-San José	-
		03	San José-Los Ángeles	-
			Los Ángeles-San José	-
		04	San José-Los Ángeles	-
			Los Ángeles-San José	-
		05	San José-Los Ángeles	-
			Los Ángeles-San José	-
		06	San José-Los Ángeles	-

			Los Ángeles-San José	-
		07	San José-Los Ángeles	-
			Los Ángeles-San José	-
		08	San José-Los Ángeles	-
			Los Ángeles-San José	-
		09	San José-Los Ángeles	-
			Los Ángeles-San José	-
		10	San José-Los Ángeles	-
			Los Ángeles-San José	-
		11	San José-Los Ángeles	-
			Los Ángeles-San José	-
		12	San José-Los Ángeles	-
			Los Ángeles-San José	-

B. Ejemplo gráficas

Ejemplo gráfica generada en Matlab para el 30 de abril de 2008 en el enlace de Chicago y protocolo TCP.



Ejemplo gráfica generada en Matlab para el 30 de abril de 2008 en el enlace de Chicago y protocolo UDP.